Hum Genet (2007) 121:93–100
DOI 10.1007/s00439-006-0277-z

ORIGINAL INVESTIGATION

Analysis of segmental duplications reveals a distinct pattern of continuation-of-synteny between human and mouse genomes

Michael R. Mehan · Maricel Almonte · Erin Slaten · Nelson B. Freimer · P. Nagesh Rao · Roel A. Ophoff

Received: 12 June 2006 / Accepted: 5 October 2006 / Published online: 8 November 2006
© Springer-Verlag 2006

Abstract About 5% of the human genome consists of large-scale duplicated segments of almost identical sequences. Segmental duplications (SDs) have been proposed to be involved in non-allelic homologous recombination leading to recurrent genomic variation and disease. It has also been suggested that these SDs are associated with syntenic rearrangements that have shaped the human genome. We have analyzed 14 members of a single family of closely related SDs in the human genome, some of which are associated with common inversion polymorphisms at chromosomes 8p23 and 4p16. Comparative analysis with the mouse genome revealed syntenic inversions for these two human polymorphic loci. In addition, 12 of the 14 SDs, while absent in the mouse genome, occur at the breaks of synteny; suggesting a non-random involvement of these sequences in genome evolution. Furthermore, we observed a syntenic familial relationship between 8 and 12 breakpoint-loci, where broken synteny that ends at one family member resumes at another, even across different chromosomes. Subsequent genome-wide assessment revealed that this relationship, which we named continuation-of-synteny, is not limited to the 8p23 family and occurs 46 times in the human genome with high frequency at specific chromosomes. Our analysis supports a non-random breakage model of genomic evolution with an active involvement of segmental duplications for specific regions of the human genome.

Electronic supplementary material Supplementary material is available in the online version of this article at http://dx.doi.org/10.1007/s00439-006-0277-z and is accessible for authorized users.

M. R. Mehan · M. Almonte · E. Slaten · N. B. Freimer · R. A. Ophoff (✉)
Center for Neurobehavioral Genetics,
Neuropsychiatric Institute,
University of California Los Angeles,
Gonda Building Room 3506,
695 Charles E. Young Drive South,
Los Angeles, CA 90095, USA
e-mail: Ophoff@ucla.edu

R. A. Ophoff
Department of Human Genetics,
University of California Los Angeles,
Gonda Building Room 3506,
695 Charles E. Young Drive South,
Los Angeles, CA 90095, USA

P. N. Rao
Molecular Cytogenetics Laboratories,
Department of Pathology,
University of California Los Angeles,
10833 Le Conte Avenue,
Los Angeles, CA 90095, USA

Introduction

Large segmental duplications (SDs) are estimated to make up about 5% of the human genome (Bailey et al. 2002; Cheung et al. 2003). Several reports indicate that SDs are associated with syntenic rearrangements in the human genome relative to mouse (Armengol et al. 2003; Bailey et al. 2004; Bi et al. 2002; Gimelli et al. 2003; Valero et al. 2000). It has been suggested that segmental duplications are driving forces for evolutionary rearrangements and variation in chromosomal structure (Armengol et al. 2003). However, the nature and pattern of SDs shaping mammalian genomes remains to be elucidated. What is clear, however, is that large duplicated segments with near identical sequences can be hotspots for the occurrence of non-allelic

homologous recombination or unequal crossing-over. These events can lead to genomic variations such as deletions, duplications, inversions, or translocations. Changes in gene dosage or rearrangements in chromosomal architectural structure mediated by SDs can lead to human disease traits and are called genomic disorders (Inoue and Lupski 2002; Lupski 1998). Recently, it has become apparent that a number of genomic variants and rearrangements occur frequently in the general population (Iafrate et al. 2004; Sebat et al. 2004; Sharp et al. 2005). Investigation of specific loci or chromosomal regions has identified a few known common variations in genomic architecture. Examples of this are found on chromosome 8p23 and 4p16 where common submicroscopical inversion polymorphisms are described (Giglio et al. 2001, 2002). Based on the structural features of SDs in opposite orientation surrounding the inversions, i.e. palindromic SDs (PSDs), we recently performed a genome-wide survey to identify additional loci in the human genome with similar higher order structure (Mehan et al. 2004). We identified more than 200 distinct loci that could contain genomic variations such as submicroscopical inversions similar to that observed on 8p23 and 4p16. Some of these have already been shown to be involved in genomic disorders and inversion polymorphisms such as Williams–Beuren syndrome (WBS, MIM 194050) at 7q11 (Osborne et al. 2001), Angelman syndrome (AS; MIM 105830) on 15q11–q13, (Gimelli et al. 2003), Emery–Dreifuss muscular dystrophy (EDMD, MIM 310300) on chromosome Xq28 (Small et al. 1997) and the microtubule-associated protein tau (MAPT) containing region at 17q21.31 (Cruts et al. 2005; Stefansson et al. 2005). The widespread distribution of these structures suggests a potential high degree of plasticity of the human genome even beyond the list of loci of already known genomic disorders. In addition to the strong sequence similarity between the 8p23 and 4p16 PSD sequences, we identified two additional PSD loci and six single SDs composed of the same duplicated sequences elsewhere in the human genome. These duplicated sequences belong to the previously identified olfactory receptor (OR)-containing 7E SD family (Newman and Trask 2003). Our study of human–mouse conserved synteny surrounding these segmental duplications led to the discovery of an association between pairs of breakpoints in human–mouse conserved synteny and their flanking sequence in the human genome. This association, which we refer to as "continuation-of-synteny", involves nearly identical segmental duplications that flank breakpoints in human–mouse synteny at two different locations in the human genome. These segmental duplications are involved in continuation-of-synteny if a syntenic block that ends at one of these breakpoints continues at the other SD containing breakpoint. To determine the extent of this phenomenon in the human genome we performed a genome-wide examination of human and mouse genomes. The results revealed that this specific pattern of syntenic rearrangement involving related segmental duplications is not uncommon in the human genome. Additionally, many of the segmental duplications involved in continuation-of-synteny form inter-related families. We discuss two of the larger family structures here: the 8p23 family, which is a subset of the OR 7E family of SDs; and three overlapping families of SDs located on chromosome 17.

Results

Genome-wide continuation-of-synteny

We conducted a genome-wide search for human genome segmental duplications involved in continuation-of-synteny with the mouse genome. Of the 283 pairs of syntenic breakpoints that we identified, 46 shared at least 5 kb of duplicated sequence with at least 85% sequence identity (Supplemental data). Many of these syntenic breakpoints contained multiple segmental duplications that clustered together, often in a consistent order. We manually mapped each of these loci to the current builds of the human and mouse genomes and found that all but two continuation-of-synteny events remained present.

The continuation-of-synteny segmental duplications we identified are primarily intra-chromosomal, with only three pairs located on different chromosomes. Mouse chromosome 11 had 11 continuation-of-synteny loci involved in segmental duplications in the human genome. This was the most for any mouse chromosome while mouse chromosomes 12, 17, and 18 contained none. We noted that human segmental duplications that are involved in continuation-of-synteny events on the same mouse chromosome often shared strong sequence identity, specifically on mouse chromosomes 5, 7, 9, and 11. Our search for continuation-of-synteny events involving segmental duplication in the mouse genome compared to the human genome did not return any SD blocks greater than 5 kb.

Inter-chromosomal 8p23 Family

In the human genome freeze of July 2003, we identified four pairs and six single occurrences of a segmental duplication that share at least 90% identity over more than 20 kb of "fuguized" sequence of the distal member of the segmental duplication pair on 8p23 ($8p23^{dist}$) (Fig. 1).

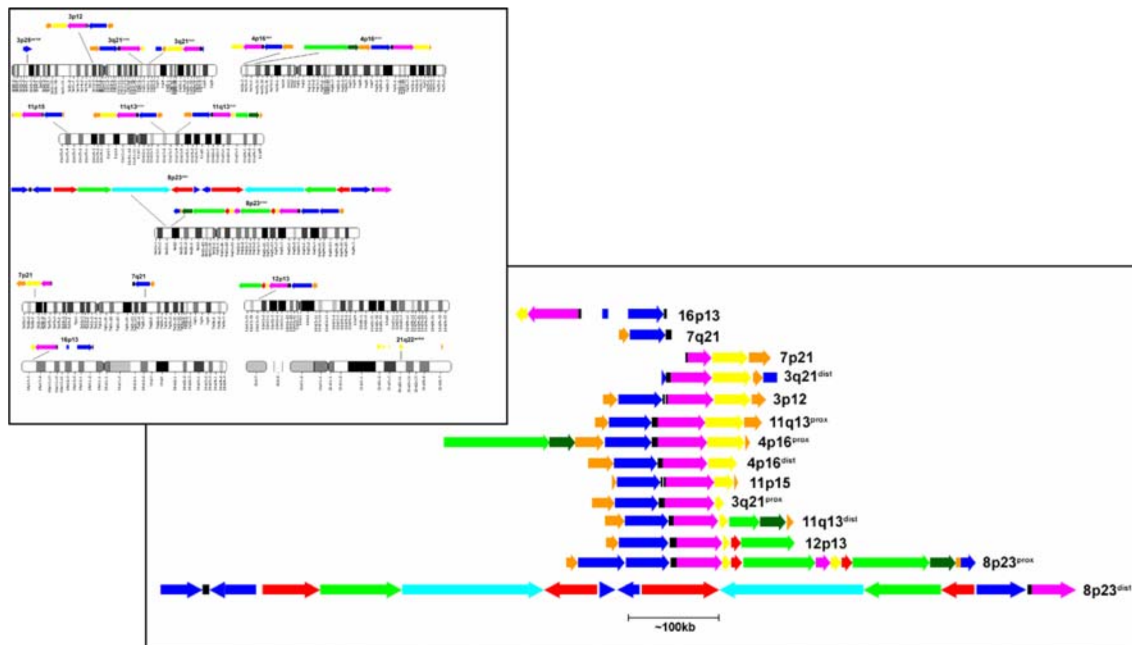

Fig. 1 The chromosomal distribution of family of 8p23 SDs. Duplicated segments that share sequence similarity are the same color. Identification of blocks was obtained by using BLAST sequence alignment tools

The "fuguization" process reduces computation time and background noise, by removing both repetitive elements masked by Repeat Masker (A.F.A. Smit and P. Green, unpublished), and non-sequenced gaps. The segmental duplications from each pair exist in opposite orientation to one another and are separated by no more than 5 Mb. The segmental duplication pairs on chromosome 8p23 and 4p16 flank known inversions, whereas the pairs on chromosomes 3q21–q22 and 11q13 do not harbor any known genomic rearrangements. Of the six unpaired segmental duplications, two are located on chromosomes that also contain segmental duplication pairs (3p21 and 11p15). Of the remaining single occurrences two are located on chromosome 7 (7p22 and 7q21), and one each on chromosome 12p13 and 16p13. Some sequence overlap was found with a region on chromosomes 3p25 and 21q, but since the sequence similarity was to a small subset (<20 kb) of the duplicated segments and very fragmented (not contiguous), it was decided not to consider these two loci in our subsequent analyses.

Of the 12 members that coincide with a break in human–mouse synteny, eight are involved in at least one continuation-of-synteny with another member of the family. Furthermore, in every case of continued synteny involving two SD loci, the orientation of the SDs is consistent. For example, if a syntenic block ends on the left hand side of the first SD, the syntenic block continues on the right hand side of the SD at the other locus (Fig. 2).

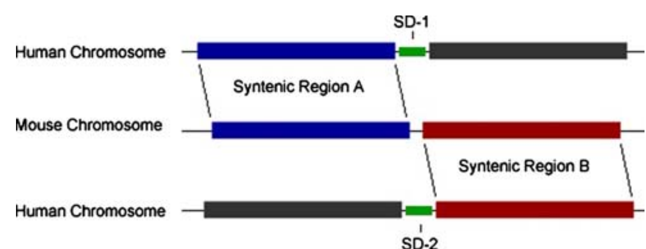

Fig. 2 Diagram representing a human continuation-of-synteny event with the mouse genome. Here two syntenic blocks located on different human chromosomes corresponds to continuous blocks in the mouse genome. These syntenic blocks are flanked by highly identical (>85% sequence identity) segmental duplications (SD-1, SD-2)

We observed that the two chromosomes that harbor known inversions, 8p23 and 4p16, show strong syntenic evidence of an inversion between the human and mouse assemblies (Fig. 3a, b). If the sequence on 8p23 between the pair (distal and proximal) of SDs is reversed, it completes the break in synteny that occurs at the $8p23^{prox}$ locus. If the sequence between the pair of SDs on chromosome 4p16 is inverted, it completes the breaks in synteny at both SD loci. We observed a similar pattern on chromosome 11, whereby if non-homologous recombination were to occur between 11p15 and $11q13^{dist}$ it would complete both breaks in synteny at both loci (Fig. 3c). We also identified a single continuation-of-synteny event involving 3p25 and 12p13.

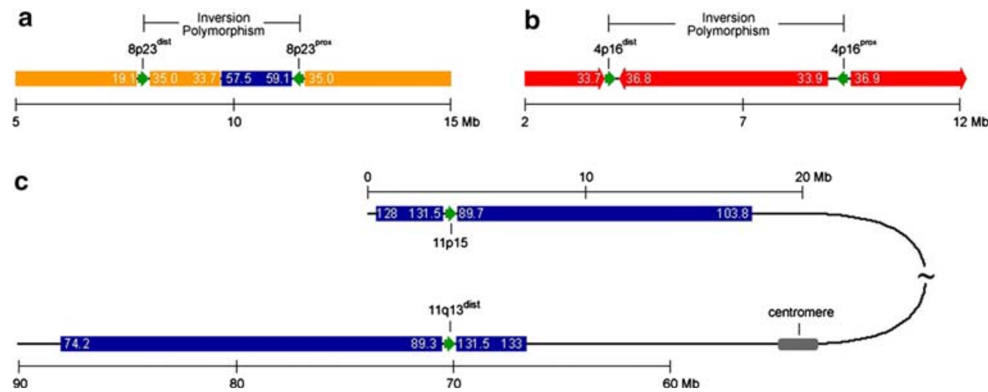

Fig. 3 Examples of human genome continuation-of-synteny. Syntenic blocks are colored according to mouse chromosome. The syntenic block location in the mouse assembly is displayed at the end of the blocks in megabases and *arrows* depict syntenic orientation. **a** One instance of continuation-of-synteny involving 8p23^{prox} and 8p23^{dist}. By reversing the sequence between the paired SDs, a larger syntenic region is formed from two mouse chromosome 8 syntenic blocks that spans the 8p23^{prox} SD. **b** Two instances of continuation-of-synteny on 4p16. Reversing the sequence between the paired SDs resolves both breaks of synteny with mouse chromosome 5. **c** Two instances of continuation of synteny involving 11p15 and 11q13^{dist} with mouse chromosome 5. Chromosome 11 is shown bent to align the orientation of the SDs

All syntenic blocks involved in continuation of synteny with in this family are derived from mouse chromosomes 5, 6, 7, and 8. Mouse chromosomes 5 and 6 are almost completely composed of sequences flanking the 8p23 SD family. The syntenic blocks flanking both 8p23 SDs from chromosome 7 and the pair from human chromosome 4 encompass 85% of mouse chromosome 5. Mouse chromosome 6 is 88% composed of syntenic blocks flanking the SDs from chromosome 7, all the SDs from chromosome 3, as well as the single SD on chromosome 12. Mouse chromosome 7 contains the sequences flanking the 11q13^{dist} and 11p15 SDs however these syntenic blocks constitute only a small part of the chromosome. Mouse chromosome 8 harbors the single continuation-of-synteny at 8p23^{prox}.

Chromosome 17 intra-chromosomal families

Our algorithm identified 11 instances of continuation of synteny involving 13 SD loci from human chromosome 17. Nine of the 13 SD loci were involved in two continuation-of-synteny events and the remaining 4 were only involved in 1. Analysis of the SD loci revealed three distinct segmental duplication blocks (Fig. 4a).

All syntenic blocks from human chromosome 17 are located on mouse chromosome 11 (Fig. 4). Previous studies comparing human–mouse syntenic block rearrangements have determined that these syntenic blocks can be reordered to match the syntenic order of mouse chromosome 11 through a series of 11 inversions (Bourque et al. 2004). These 11 inversions occur at 15 breakpoints, 5 of which are used more than once. We have determined that each of the breakpoints from the proposed inversions share duplicated sequence, and these sequences are members of the three families we identified. In one case of breakpoint reuse, the pair of similar SDs that flank the first proposed inversion are from a different family than the pair that flank second proposed inversion. This is due to neighboring SDs at the same locus (Fig. 4).

Discussion

We have discovered a relationship between breakpoints in human–mouse synteny in particular regions of the human genome that strongly supports a non-random breakage model of genomic evolution. We investigated the extent of this relationship in the human genome and identified 46 pairs of segmental duplications with a pattern of continuation-of-synteny between mouse and human genomes involving sequences immediately flanking these SDs. Our analysis revealed the SDs involved in this continuation-of-synteny phenomenon cluster into SD families as seen in the 8p23 and chromosome 17 families. These families contain segmental duplications that flank known inversion polymorphisms in the human genome, and may be instrumental in shaping the human genome.

Continuation-of-synteny involving the 8p23 SD family

Continuation of synteny provides strong support for the non-random breakage model of chromosomal evolution. Previous studies have shown that breakpoints in synteny are enriched with segmental duplications but hypothesized that this association did not necessarily

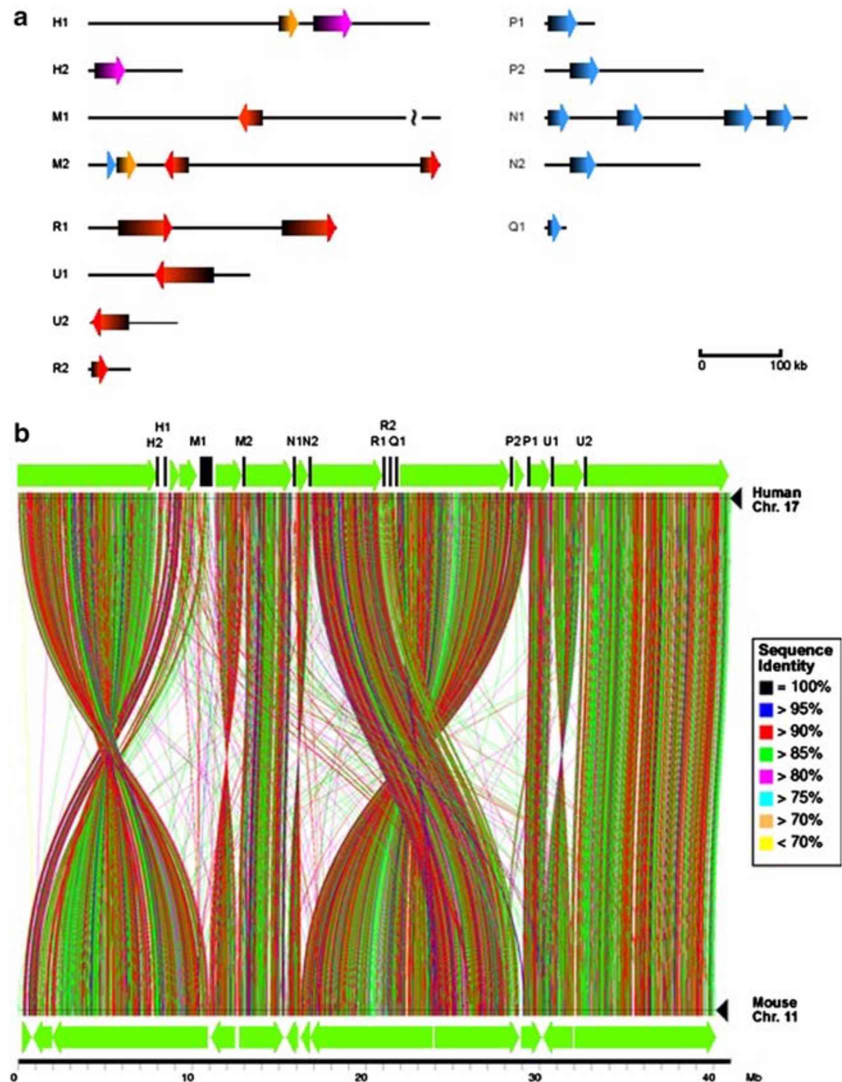

Fig. 4 Segmental duplications and surrounding sytenic blocks between human chromosome 17 and the distal-q arm of mouse chromosome 11. Both chromosomes and the segmental duplications have been fuguized. **a** The 13 SD loci identified by our algorithm that are involved in at least one continuation-of-synteny event. The *colored arrows* of the same color represent segmental duplication blocks that share sequence identity. The names of the loci correspond to pairs identified by our algorithm. **b** A visual representation of BLAST alignments between human chromosome 17 and mouse chromosome 11 illustrating the continuation-of-synteny phenomenon. The *black vertical bars* above the human sequence represent the location of human-specific segmental duplications and the *green arrows* indicate syntenic block orientation

imply causality (Armengol et al. 2003; Bailey et al. 2004). While none of the sequences of the 8p23 SD family were present in the mouse genome, we noted that 12 of the 14 SD loci in the human genome coincided with a break of synteny in the mouse. This observation confirms earlier reports describing duplicated segments to be associated with syntenic rearrangements (Armengol et al. 2003; Bailey et al. 2004). Although none of the duplicated sequence is located in the mouse genome, a striking observation is that mouse chromosomes 5 and 6 are almost entirely composed of sequence flanking the 8p23 family. One possible explanation for this is that the 8p23 SD family mediated both inter and intra-chromosomal syntenic rearrangement events between the ancestral mammalian chromosomes, and especially within the primate lineage.

The involvement of the 8p23 SD family in syntenic rearrangements in the primate lineage is further supported by one member of the 8p23 SD family located at 11q13. The proximal 11q13 duplication is not located at a break of human–mouse synteny and absent from the chimpanzee assembly, suggesting that it is a human-specific duplication. Our experimental in situ hybridization analyses of this region were suggestive for the absence of the 11q13^{prox} duplication in the chimpanzee genome as well (data not shown). Previous cytogenetic work of Trask and colleagues (Trask et al. 1998), studying the distribution of olfactory receptor gene family clusters (also part of the 8p23 SD family) in different primates species showed a decrease of probe intensity for human 11q13 in the orthologous chimpanzee region, but did not discuss the presence of only one copy. Further detailed molecular analysis is required to confirm the absence of the 11q13^{prox} duplicated segments in the chimpanzee genome, and show that this segmental duplication event is unique to the lineage of *Homo sapiens*, and therefore not likely to be

involved in syntenic rearrangements separating rodent and primate genomes. It is possible that a single instance of the continuation-of-synteny or the lack thereof (i.e. on 11q13^{prox}) results from an artifact in the genome assembly but that does not change the overall picture of specific patterns of synteny in the human genome.

Comparative analysis as predictive tool for inversion polymorphisms

In a recent study comparing the genomes of human and chimpanzee, a large number of syntenic inversions were identified, some of which were later shown to be human inversion polymorphisms (Feuk et al. 2005). Our observations support this approach even with comparative analysis between two species that are as distant as human and mouse. In many known common inversion polymorphisms, such as the 4p16 and 8p23 inversion regions, the human–mouse synteny in these regions was reversed and flanked by one or two continuation of synteny loci. A similar finding of syntenic inversions between mouse and human orthologous sequences was previously reported for the WBS region on 7q11 (DeSilva et al. 2002; Valero et al. 2000) and AS region on chromosome 15q11–q13 (Gimelli et al. 2003). These two regions are not only known to be involved in chromosomal rearrangements leading to genomic disorders but also to harbor common inversion polymorphisms (Gimelli et al. 2003; Osborne et al. 2001). The recently identified inversion polymorphism at 17q21.31 containing the MAPT gene (Cruts et al. 2005; Stefansson et al. 2005) shows similar syntenic features in a comparison between human and mouse genomes. These multiple observations of human–mouse syntenic inversions of human loci together with known genomic inversion polymorphisms suggest that comparative genomic analysis may be a powerful tool for predicting the location of submicroscopical inversion polymorphisms and perhaps other types of genomic variations. The potential for this method to predict the involvement of the paralogous segments in genomic variation is exemplified by the identification of PSD structures at 3q21–q22 and 16q22.

Implication for genomic evolution

The SDs that flank known common inversions polymorphisms in the human genome are consistently in opposite ("palindromic") orientation indicating that SDs role in mediating inversions is orientation dependent (Mehan et al. 2004). This palindromic pair structure is seen in four instances within the 8p23 SD family. We further observed that all 8p23 family SDs that are part of continuation-of-synteny events, regardless of their proximity, exhibit the same orientation relative to their flanking syntenic blocks, This is true for continuation-of-synteny events with paired SDs (e.g. 4p16), non-paired on the same chromosome (e.g. 11q13^{dist} and 11p15), as well as for inter-chromosomal instances of continuation of synteny (12p13 and 3p25). This suggests first that this family of SDs is not only involved in more recent common inversion polymorphisms, but may also have played a role in evolutionary primate chromosomal rearrangements. The fact that we did not detect any mouse–human continuation-of-synteny patterns in the mouse genome suggests that this phenomenon may have occurred only in specific lineages. Alternatively, the absence of these events in the mouse genome may be caused by an incomplete genome assembly in which segmental duplications are underrepresented.

Our genome-wide search revealed that approximately 17% of breakpoints in human–mouse synteny contain at least one continuation of synteny event. Recently, Bourque et al. (2004) described a much higher rate of rearrangements in rodents than in the human lineage, implicating the existence of rearrangement hotspots. In a three-way comparison of human chromosome 17 with mouse chromosome 11 and rat chromosome 10 (h17/m11/r10) they reconstructed the genomic architecture and assigned rearrangement events to the particular lineages. Interestingly, each of the human chromosome 17 rearrangement hotspots involved in syntenic inversions described by Bourque et al. (2004) contain related duplicated segments that show continuation-of-synteny in our analysis. Our finding implies an active involvement of SDs in syntenic rearrangements in the primate lineage for this part of the genome, expanding upon Bourque's proposed model of breakpoint re-usage. Recent discussions about human genome evolution have favored a non-random model, in which a tendency is observed for chromosomal breaks to reoccur at certain "fragile" genome regions (Armengol et al. 2003; Bailey et al. 2004; Pevzner and Tesler 2003). This model has been contrasted with a random breakage model of genomic evolution postulated by Ohno more than three decades ago (Ohno 1973) and supported by others (Nadeau and Taylor 1984). Our observation of human/mouse continuation-of-synteny affecting specific regions of these genomes strongly supports the more recent hypothesis of non-random model of breaks of synteny. However, our findings also suggest a possible "local" effect of this phenomenon, which may indicate that both models may be valid for different parts of the human genome.

Further systematic analysis of regions in the human genome considering duplicated segments as well as syntenic rearrangements in genomes of other species is required and should provide more insight into mechanisms of genome evolution, plasticity, and possibly common genomic variation.

Methods

Identification of 8p23 segmental duplication family

To determine the extent of genome wide duplication of the segmental duplications on chromosome 8p23, we used UCSC BLAT search to align sequence from the distal segmental duplication ($8p23^{dist}$) against the July 2003 Freeze (hg16) of the human genome. Our criterion for a duplicated locus was at least 90% of sequence identity over 20 kb. All sequence analyses were performed after "fuguization"(Bailey et al. 2001), i.e. removal of repetitive elements masked by Repeat Masker (A.F.A. Smit and P. Green, unpublished), and unsequenced gaps in the genome assembly.

Genome-wide search for continuation-of-synteny

Analysis of genome-wide continuation of synteny was conducted using the Mouse Alignment Net data set (netMm4 Position Table) from the UCSC Genome Browser. The Mouse Alignment Net data set contains consensus alignment chains derived from blastz alignments of human build 33 and mouse build mm4. We "fuguized" the human and mouse genomes by removing repetitive sequences and only considered alignment chains greater than 200 kb to exclude the spurious alignment chains that often localize in duplicated regions.

Human continuation-of-synteny is defined as a relationship between two distinct breakpoints in human–mouse synteny. The criteria of continuation-of-synteny are the following: (1) a continuous syntenic sequence in the mouse genome is broken into two non-consecutive locations in the human genome, and (2) the two distinct breakpoints in the human genome are flanked by at least 5 kb of segmental duplications with at least 85% identity.

To determine all instances that satisfy the first criterion we sorted the alignment chains by mouse chromosomal position. For each pair of consecutive alignment chains in the mouse genome, we verified that they aligned to two non-consecutive regions in the human genome. Then to determine whether the breaks in synteny shared flanking identical segmental duplications, the sequence flanking both breaks in conserved synteny was aligned using BLAST. The flanking sequence for each break in conserved human–mouse synteny ended at the beginning of the next consecutive alignment chain in the human genome. If these two regions between blocks of conserved synteny in the human genome contained alignments that exceeded 5 kb in total length with 85% sequence identity we considered this a single instance of continuation of synteny. Inversions polymorphisms that produce reciprocal examples of continuation of synteny, like the 4p16 inversion for example were considered only one instance of continuation of synteny.

Acknowledgments The study was supported by grants from NIGMS (GM068875) and NARSAD Young Investigator Award (to RAO).

References

Armengol L, Pujana MA, Cheung J, Scherer SW, Estivill X (2003) Enrichment of segmental duplications in regions of breaks of synteny between the human and mouse genomes suggest their involvement in evolutionary rearrangements. Hum Mol Genet 12:2201–2208

Bailey JA, Baertsch R, Kent WJ, Haussler D, Eichler EE (2004) Hotspots of mammalian chromosomal evolution. Genome Biol 5:R23

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE (2002) Recent segmental duplications in the human genome. Science 297:1003–1007

Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE (2001) Segmental duplications: organization and impact within the current human genome project assembly. Genome Res 11:1005–1017

Bi W, Yan J, Stankiewicz P, Park SS, Walz K, Boerkoel CF, Potocki L, Shaffer LG, Devriendt K, Nowaczyk MJ, Inoue K, Lupski JR (2002) Genes in a refined Smith–Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. Genome Res 12:713–728

Bourque G, Pevzner PA, Tesler G (2004) Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. Genome Res 14:507–516

Cheung J, Estivill X, Khaja R, MacDonald JR, Lau K, Tsui LC, Scherer SW (2003) Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. Genome Biol 4:R25

Cruts M, Rademakers R, Gijselinck I, van der Zee J, Dermaut B, de Pooter T, de Rijk P, Del-Favero J, van Broeckhoven C (2005) Genomic architecture of human 17q21 linked to frontotemporal dementia uncovers a highly homologous family of low-copy repeats in the tau region. Hum Mol Genet 14:1753–1762

DeSilva U, Elnitski L, Idol JR, Doyle JL, Gan W, Thomas JW, Schwartz S, Dietrich NL, Beckstrom-Sternberg SM, McDowell JC, Blakesley RW, Bouffard GG, Thomas PJ, Touchman JW, Miller W, Green ED (2002) Generation and comparative analysis of approximately 3.3 Mb of mouse genomic sequence orthologous to the region of human chromosome 7q11.23 implicated in Williams syndrome. Genome Res 12:3–15

Feuk L, MacDonald JR, Tang T, Carson AR, Li M, Rao G, Khaja R, Scherer SW (2005) Discovery of human inversion polymorphisms by comparative analysis of human and chimpanzee DNA sequence assemblies. PLoS Genet 1:e56

Giglio S, Broman KW, Matsumoto N, Calvari V, Gimelli G, Neumann T, Ohashi H, Voullaire L, Larizza D, Giorda R, Weber JL, Ledbetter DH, Zuffardi O (2001) Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. Am J Hum Genet 68:874–883

Giglio S, Calvari V, Gregato G, Gimelli G, Camanini S, Giorda R, Ragusa A, Guerneri S, Selicorni A, Stumm M, Tonnies H, Ventura M, Zollino M, Neri G, Barber J, Wieczorek D, Rocchi M, Zuffardi O (2002) Heterozygous submicroscopic inversions involving olfactory receptor-gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. Am J Hum Genet 71:276–285

Gimelli G, Pujana MA, Patricelli MG, Russo S, Giardino D, Larizza L, Cheung J, Armengol L, Schinzel A, Estivill X, Zuffardi O (2003) Genomic inversions of human chromosome 15q11-q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. Hum Mol Genet 12:849–858

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C (2004) Detection of large-scale variation in the human genome. Nat Genet 36:949–951

Inoue K, Lupski JR (2002) Molecular mechanisms for genomic disorders. Annu Rev Genomics Hum Genet 3:199–242

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet 14:417–422

Mehan MR, Freimer NB, Ophoff RA (2004) A genome-wide survey of segmental duplications that mediate common human genetic variation of chromosomal architecture. Hum Genomics 1:335–344

Nadeau JH, Taylor BA (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. Proc Natl Acad Sci USA 81:814–818

Newman T, Trask BJ (2003) Complex evolution of 7E olfactory receptor genes in segmental duplications. Genome Res 13(5):781–793

Ohno S (1973) Ancient linkage groups and frozen accidents. Nature 244:259–262

Osborne LR, Li M, Pober B, Chitayat D, Bodurtha J, Mandel A, Costa T, Grebe T, Cox S, Tsui LC, Scherer SW (2001) A 1.5 million-base pair inversion polymorphism in families with Williams–Beuren syndrome. Nat Genet 29:321–325

Pevzner P, Tesler G (2003) Human and mouse genomic sequences reveal extensive breakpoint reuse in mammalian evolution. Proc Natl Acad Sci USA 100:7672–7677

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M (2004) Large-scale copy number polymorphism in the human genome. Science 305:525–528

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE (2005) Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77:78–88

Small K, Iber J, Warren ST (1997) Emerin deletion reveals a common X-chromosome inversion mediated by inverted repeats. Nat Genet 16:96–99

Stefansson H, Helgason A, Thorleifsson G, Steinthorsdottir V, Masson G, Barnard J, Baker A, Jonasdottir A, Ingason A, Gudnadottir VG, Desnica N, Hicks A, Gylfason A, Gudbjartsson DF, Jonsdottir GM, Sainz J, Agnarsson K, Birgisdottir B, Ghosh S, Olafsdottir A, Cazier JB, Kristjansson K, Frigge ML, Thorgeirsson TE, Gulcher JR, Kong A, Stefansson K (2005) A common inversion under selection in Europeans. Nat Genet 37:129–137

Trask BJ, Massa H, Brand-Arpon V, Chan K, Friedman C, Nguyen OT, Eichler E, van den Engh G, Rouquier S, Shizuya H, Giorgi D (1998) Large multi-chromosomal duplications encompass many members of the olfactory receptor gene family in the human genome. Hum Mol Genet 7:2007–2020

Valero MC, de Luis O, Cruces J, Perez Jurado LA (2000) Fine-scale comparative mapping of the human 7q11.23 region and the orthologous region on mouse chromosome 5G: the low-copy repeats that flank the Williams–Beuren syndrome deletion arose at breakpoint sites of an evolutionary inversion(s). Genomics 69:1–13